

**IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
WACO DIVISION**

XOCKETS, INC.,

Plaintiff,

v.

**NVIDIA CORPORATION,
MICROSOFT CORPORATION, and
RPX CORPORATION,**

Defendants.

Civil Action No. 6:24-cv-453

JURY TRIAL DEMANDED

DECLARATION OF JASON SHEASBY

I, Jason Sheasby, declare as follows:

1. I am a partner at Irell & Manella LLP and represent Xockets, Inc. (“Xockets”) in this matter. I submit this declaration in support of Xockets’ motion for a preliminary injunction. Unless otherwise specified, I have personal knowledge of the facts stated herein, and, if called as a witness, could and would testify to such facts competently under oath.

NVIDIA and Microsoft’s Admissions Regarding Their Use of Xockets’ DPU Architecture

2. NVIDIA Corporation (“NVIDIA”) and Microsoft Corporation (“Microsoft”) have made numerous admission regarding their use and the importance of DPU architecture enabled by Xockets’ technology. These include the following:

3. NVIDIA publicly states that “[t]he best definition of the DPU’s mission is to offload, accelerate, and isolate infrastructure workloads” and further explains each function¹:

¹ Ex. 1 (<https://developer.nvidia.com/blog/offloading-and-isolating-data-center-workloads-with-bluefield-dpu/>).

- **Offload:** Take over infrastructure tasks from the server CPU so more CPU power can be used to run applications.
- **Accelerate:** Run infrastructure functions more quickly than the CPU can, using hardware acceleration in the DPU silicon.
- **Isolate:** Move key data plane and control plane functions to a separate domain on the DPU, both to relieve the server CPU from the work and to protect the functions in case the CPU or its software are compromised.

A DPU should be able to do all three tasks.

4. NVIDIA's CEO, Jensen Huang, explained the significance of DPU technology:

[W]hen you take a large scale problem that spans the whole datacenter – it doesn't fit in a single computer – and you accelerate the computation by several orders of magnitude... then the network becomes the problem, and it needs to be very fast. And so that's the reason our relationship with Mellanox goes back a decade and we've been working with them for quite a long time. The networking problem is much, much more complex than just having faster and faster networking. And the reason for that is because of the amount of data that you are transmitting, synchronizing, collecting, and reducing across this distributed data center-scale computer and the computation on the fabric itself is complicated. ... Putting intelligence in the network – and processing in the network – is vitaly important to performance.²

5. Huang also described the importance of offloading to a DPU, just as disclosed in the Xockets' Patents:

A lot of the data movement is done on the CPU. It makes no sense. *You have to offload that to a data processing unit, or DPU, which is what a SmartNIC is.* A lot of datacenters today have every single packet that is transmitted secured because you want to reduce the attack surface of the datacenter until it's basically every single transaction. There's no way you going to do that on the CPU. *So you have to move the networking stack off. You want to move the security stack off and you want to move the data processing and data movement stack off.* And this is something that you want to do right at the NIC before it even comes into the computer and at the

² Ex. 2 (<https://www.nextplatform.com/2020/04/27/nvidia-plus-mellanox-talking-datacenter-architecture-with-jensen-huang/>).

NIC before it leaves the computer. The onion, celery, and carrots – you know, *the holy trinity of computing* soup – is the CPU, the GPU, and the DPU. . . . *A DPU is going to be programmable, it's going to do all of that processing that you and I have already talked about, and it's going to offload the movement of data into the granular processing of the data as it's being transmitted and keep it from ever bothering the CPUs and GPUs* and avoid redundant copies of data. *That's the architecture of the future.* And that's the reason why we're so excited about Mellanox.³

6. Since the launch of its systems that use Xockets' technology, NVIDIA's market capitalization has surged over ten-fold, from approximately \$170 billion to over \$3 trillion as of the filing of this Complaint.

7. Further explaining the importance of Xockets' revolutionary DPU architecture, Huang declared:

One of the most important things to disaggregate out of the server node and its CPU is the data processing. That is a giant amount of unnecessary CPU cores running unnecessary software in the datacenter. I don't know how much – its maybe 30 percent to 50 percent. . . . *I really do think that when you offload [to] the data processing on the SmartNIC*, when you're able to disaggregate the converged server, *when you can put accelerators anywhere in datacenter* and then can compose and reconfigure that datacenter for this specific workload – *that's a revolution.*⁴

8. NVIDIA explained that “DPUs are an essential element of modern and secure data centers in which CPUs, GPUs and DPUs are able to combine into a single computing unit that's fully programmable, AI-enabled and can deliver levels of security and compute power not previously possible.”⁵ Similarly, in an April 2021 press release, NVIDIA stated that “[a] new type

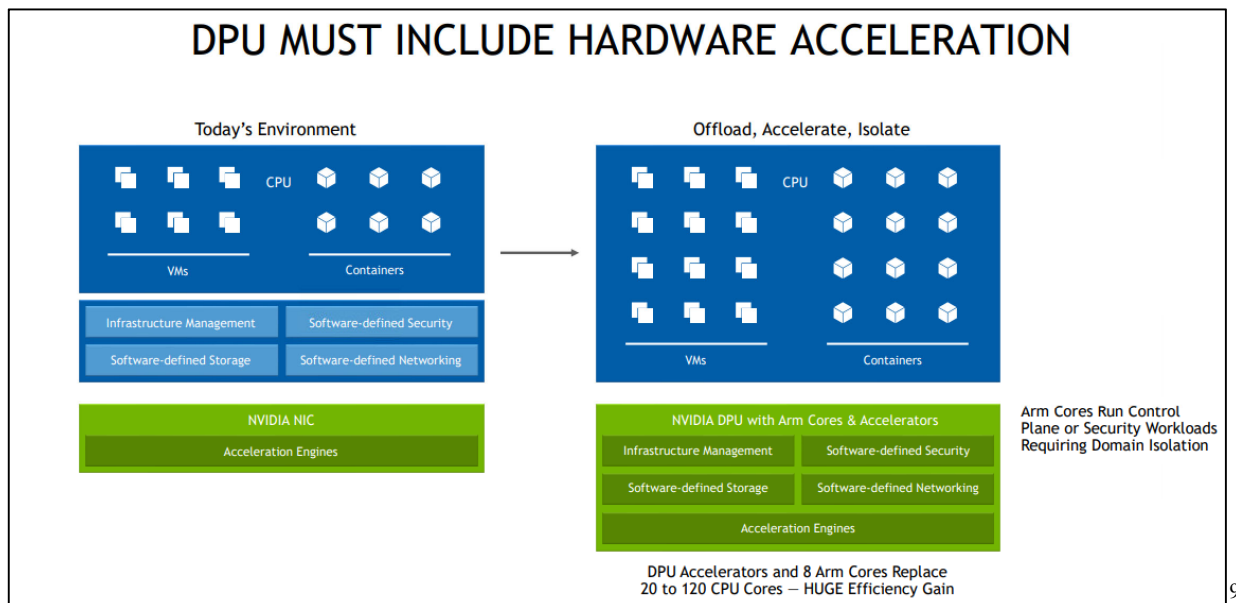
³ Ex. 2 (<https://www.nextplatform.com/2020/04/27/nvidia-plus-mellanox-talking-datacenter-architecture-with-jensen-huang/>).

⁴ Ex. 2 (<https://www.nextplatform.com/2020/04/27/nvidia-plus-mellanox-talking-datacenter-architecture-with-jensen-huang/>).

⁵ Ex. 3 (<https://nvidianews.nvidia.com/news/nvidia-introduces-new-family-of-bluefield-dpus-to-bring-breakthrough-networking-storage-and-security-performance-to-every-data-center>).

of processor, designed to process data center infrastructure software, is needed to offload and accelerate the tremendous compute load of virtualization, networking, storage, security and other cloud-native AI services. The time for BlueField DPU has come.”⁶ NVIDIA has described the DPU as a “new pillar” that is “designed to offload, accelerate, and isolate infrastructure workloads and bring efficiency and security to software defined workloads such as networking security and storage while freeing CPU resources by up to 30%.”⁷

9. In fact, NVIDIA illustrates how “[o]ffloading infrastructure tasks to the DPU improves server performance, efficiency, and security”⁸ using Xockets’ DPU computing architecture:



⁶ Ex. 4 (<https://nvidianews.nvidia.com/news/nvidia-extends-data-center-infrastructure-processing-roadmap-with-bluefield-3>).

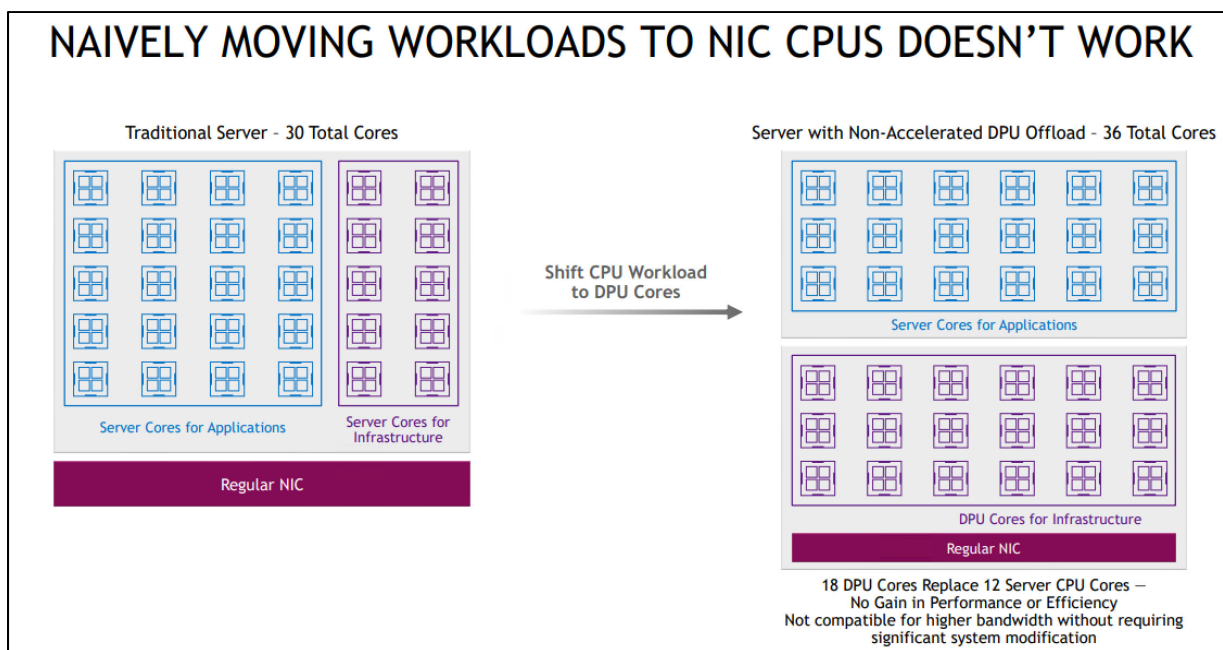
⁷ <https://www.youtube.com/watch?v=htR19dBicA>.

⁸ Ex. 1 (<https://developer.nvidia.com/blog/offloading-and-isolating-data-center-workloads-with-bluefield-dpu/>).

⁹ Ex. 5 (<https://hc33.hotchips.org/assets/program/conference/day1/HC2021.NVIDIA.IdanBurststein.v08.norecording.pdf>).

10. NVIDIA calls this a “fundamental new architecture.”¹⁰ Huang refers to the DPU paradigm as a “fundamental transition” necessitated by the fact that “CPU scaling [Moore’s law] has ended. We need a new computing approach and accelerated computing is the path forward. . . This way of doing computation is a reinvention from the ground up.”¹¹

11. Further, NVIDIA illustrates that Xockets’ DPU computing architecture is a breakthrough innovation, admitting that “naively moving workloads to NIC CPUs doesn’t work”¹²:



12. In April 2020, Huang appeared on CNBC to “[take] a victory lap after his company completed its long-awaited merger of chip producer Mellanox Technologies”—a \$7 billion deal:

¹⁰ Ex. 4 (<https://nvidianews.nvidia.com/news/nvidia-extends-data-center-infrastructure-processing-roadmap-with-bluefield-3>) (“Modern hyperscale clouds are driving a fundamental new architecture for data centers,” said Jensen Huang, founder and CEO of NVIDIA.”).

¹¹ Ex. 6 (https://www.nvidia.com/en-us/events/computex/?nvid=nv-int-cwmfg-130532#cid=cmptx23e_nv-int-cwmfg_en-us); see also <https://www.youtube.com/watch?v=i-wpzS9ZsCs&t=875s> (14:46-15:47).

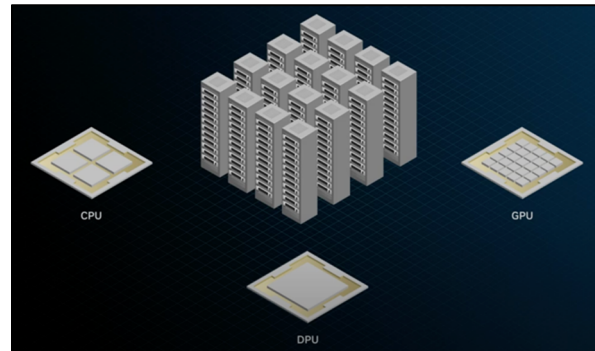
¹² Ex. 5 (<https://hc33.hotchips.org/assets/program/conference/day1/HC2021.NVIDIA.IdanBurststein.v08.norecording.pdf>).

“Man I’ve been dreaming about this. You know the most important computer today is the data center, it is the epicenter of the computer industry. And *the most important applications that run in the data center today are AI applications and Big Data analytics applications. Doing computation on artificial intelligence . . . and moving huge amounts of data around is what drives the data center architectures today.* And so we are combining the leaders of AI computing and high speed networking and data processing into one company. This is really quite extraordinary.”¹³

13. Describing NVIDIA’s adoption of the DPU computing architecture protected by Xockets’ patents, NVIDIA stated:

“Data centers are evolving and expanding to include another pillar along the CPU and GPU. *The new pillar is the Data Processing Unit or DPU.*

The NVIDIA BlueField DPU is designed to *offload, accelerate, and isolate* infrastructure workloads and bring efficiency and security to software defined [data-intensive] workloads such as networking security and storage while freeing CPU resource by up to 30%.”¹⁴



14. NVIDIA, including its CEO Huang, agrees that DPUs use a fundamental new computing architecture:

A DPU is going to be programmable . . . and it’s going to *offload the movement of data into the granular processing of the data as it’s being transmitted and keep it from ever bothering the CPUs and GPUs* and avoid redundant copies of data. *That’s the architecture of the future. And that’s the reason why we’re so excited about Mellanox.*¹⁵

¹³ Ex. 7 (<https://www.cnbc.com/2020/04/27/nvidia-ceo-calls-mellanox-acquisition-a-homerun-deal.html>).

¹⁴ NVIDIA Video illustrating the operation of its BlueField / ConnectX DPUs, (<https://www.youtube.com/watch?v=htR19rdBicA&t=3s>) (0:03–0:30).

¹⁵ Ex. 2 (<https://www.nextplatform.com/2020/04/27/nvidia-plus-mellanox-talking-datacenter-architecture-with-jensen-huang/>).

15. In fact, according to Huang, the architecture is revolutionary:

*I really do think that when you offload [to] the data processing on the SmartNIC [DPU], when you're able to disaggregate the converged server, when you can put accelerators anywhere in datacenter and then can compose and reconfigure that datacenter for this specific workload – **that's a revolution.***¹⁶

16. NVIDIA also lauds this technology as pioneering—a “reinvention from the ground up”:

“There are two fundamental transitions happening in the computer industry today.

The first trend is because CPU scaling [Moore's Law] has ended. . . . [W]e need a new computing approach and accelerated computing is the path forward.

It happened at exactly the time when a new way of doing software was discovered, deep learning [ML/AI], these two events came together and it's driving computing today: *Accelerated computing and generative AI. . . . This way of doing computation is a reinvention from the ground up.*”¹⁷



17. NVIDIA has also repeatedly recognized the importance of DPUs for cloud offload processing, and described the benefits achieved by implementing Xockets' patented inventions.

18. For example, introducing the NVIDIA BlueField-2 DPU at NVIDIA's GTC 2020 Keynote on October 5, 2020, Huang said:

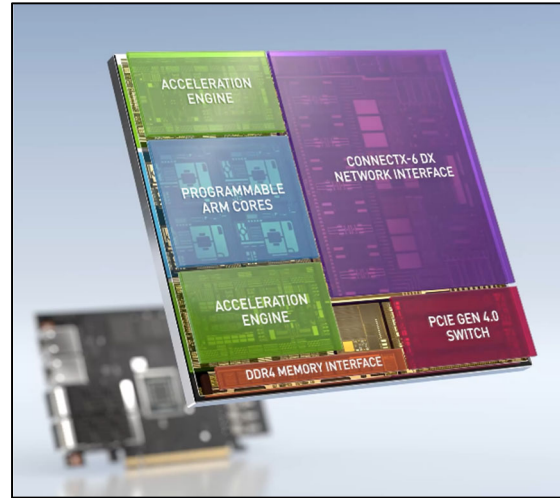
¹⁶ Ex. 2 (<https://www.nextplatform.com/2020/04/27/nvidia-plus-mellanox-talking-datacenter-architecture-with-jensen-huang/>).

¹⁷ NVIDIA CEO Jensen Huang's Keynote at Computex – Part 1 - May 2023, <https://www.youtube.com/watch?v=i-wpzS9ZsCs&t=874s> (14:34–15:46).

“Today we are announcing the BlueField-2 DPU.

It is a *programmable processor with accelerators and engines for at-line-speed processing for networking, storage, and security*. The BlueField DPU is a data center infrastructure on a chip. BlueField-2 has Arm CPUs and a whole host of state-of-the-art accelerators and hardware engines. BlueField-2 does the security processing for private, public, and hybrid clouds. . . .

BlueField-2 is a 7 billion transistor marvel. A programmable data center on-a-chip. One that we intend to support for as long as we shall live.”¹⁸



19. NVIDIA further stated that its DPUs “enable[] breakthrough networking, storage and security performance”—as claimed in the New Cloud Processor Patents.¹⁹

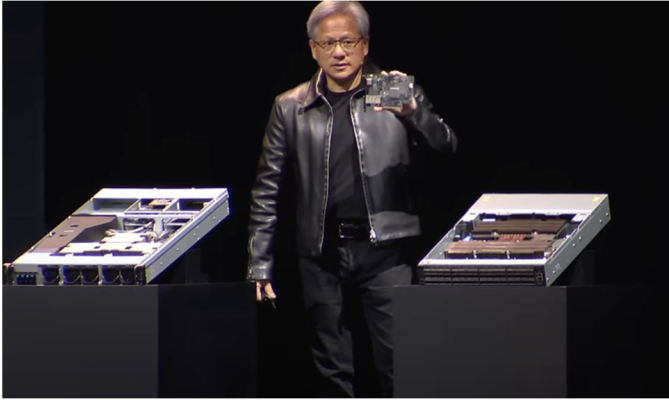
20. And NVIDIA touts the ability to “transform the data center with NVIDIA DPUs” by “offloading, accelerating, and isolating a broad range of advanced networking, storage, and security services.”²⁰

21. NVIDIA further characterized its infringing DPU-enabled architecture—first described and claimed by Xockets—as essential for high performance networks necessary for generative AI:

¹⁸ NVIDIA CEO Jensen Huang’s Keynote at GTC – Part 5 – October 2020, <https://youtu.be/MRdJ78dWjn4?t=138> (2:18–4:15).

¹⁹ Ex. 8 (<https://nvidianews.nvidia.com/news/nvidia-introduces-new-family-of-bluefield-dpus-to-bring-breakthrough-networking-storage-and-security-performance-to-every-data-center>).

²⁰ Ex. 9 (<https://www.advancedhpc.com/pages/nvidia-bluefield-data-processing-units>).



“Our new Ethernet system for AI is the Spectrum-4 switch and *the BlueField-3 SmartNIC or DPU* . . . *This is what it takes to build a high performance network. And we’re gonna take this capability to the world’s CSPs.*

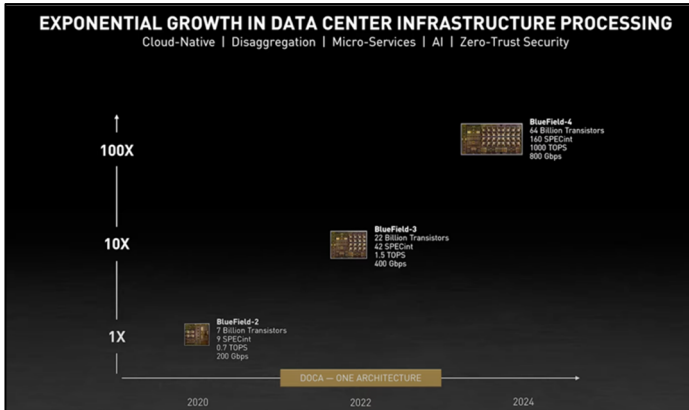
The reception has been incredible, and the reason for that is, of course, every CSP, every data center would like to turn every single data center into a generative AI data center.”²¹

22. As stated on NVIDIA’s website, the BlueField DPU “*ignites unprecedented innovation for modern data centers and supercomputing clusters.* With its robust compute power and *integrated software-defined hardware accelerators for networking, storage, and security,* BlueField creates a secure and accelerated infrastructure for any workload in any environment, *ushering in a new era of accelerated computing and AI.*”²²

23. Huang explained this further during his keynote address at the 2021 GPU Technology Conference:

²¹ NVIDIA CEO Jensen Huang’s Keynote at Computex - Part 2 - May 2023, <https://www.youtube.com/watch?v=i-wpzS9ZsCs&t=4879s> (1:21:19–1:22:21).

²² Ex. 10 (<https://www.nvidia.com/en-in/networking/products/data-processing-unit/>).

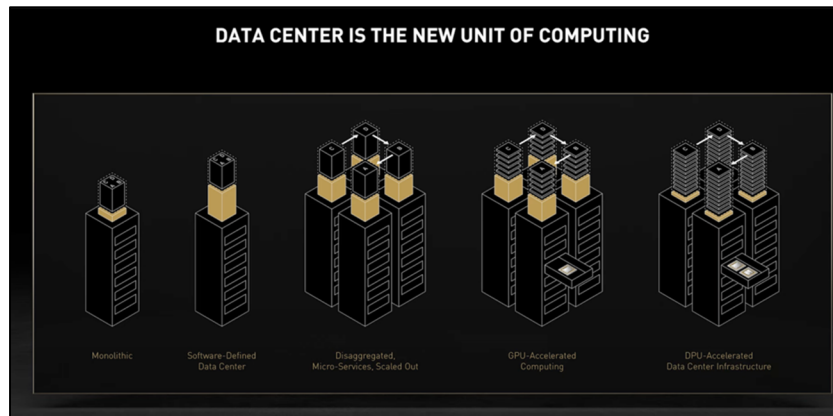


“A simple way to think about this is that *one-third of the roughly 30 million data center servers shipped each year are consumed running the software-defined data center stack.* This workload is increasing much faster than Moore’s law.

We know this because of the amount of data we are producing and moving around. *So, unless we offload and accelerate this workload, data centers will have fewer and fewer CPUs to run applications. The time for BlueField has come.*”²³

24. Huang continued:

“*Data center is the new unit of computing. Cloud computing and AI are driving fundamental changes in the architecture of data centers.* Traditionally, enterprise data centers ran monolithic software packages. Virtualization started the trend toward software-defined data centers – allowing applications to move about . . . With virtualization, *the compute, networking, storage, and security functions are emulated on software running on the CPU.* Though easier to manage, *the added CPU load reduced the data center’s capacity to run applications,* which is its primary purpose.



This illustration shows the added CPU load in the gold-colored part of the stack. Cloud computing re-architected data centers again,

²³ NVIDIA CEO Jensen Huang’s Keynote at GTC - April 2021, <https://www.youtube.com/watch?v=rzdBHBx3eJk&t=295s> (5:20–5:47).

now to provision services for billions of consumers. Monolithic applications were disaggregated into smaller microservices that can take advantage of any idle resource. . . . Data center networks became swamped by east-west traffic generated by disaggregated microservices. . . . Then, deep learning emerged. . . . Deep learning is compute-intensive, which drove adoption of GPUs. . . . *Meanwhile, the mountain of infrastructure software continues to grow. . . . The answer is a new type of chip for data center infrastructure processing like NVIDIA's BlueField DPU.*"²⁴

25. NVIDIA also conceded that DPUs deliver extraordinary TCO Savings in training large models. As NVIDIA has touted: "*A single BlueField-2 DPU can deliver the same data center services that could consume up to 125 CPU cores.* This frees up valuable CPU cores to run a wide range of other enterprise applications."²⁵

26. Elsewhere, NVIDIA has also admitted that "*[o]ne Bluefield-3 DPU delivers the equivalent data center services of up to 300 CPU cores,* freeing up valuable CPU cycles to run business-critical applications."²⁶

27. On August 30, 2022, Huang further publicly touted NVIDIA's DPUs, and explained:

"The return on investment — the benefits that DPU-enabled vSphere 8 with NVIDIA Bluefield deliver — will be so fast because it frees up so many resources for computing that the payback is going to be instantaneous It's going to be a really fantastic return."²⁷

²⁴ NVIDIA CEO Jensen Huang's Keynote at GTC - April 2021, <https://www.youtube.com/watch?v=rzdBHBx3eJk&t=4s> (0:04–1:59).

²⁵ Ex. 8 (<https://nvidianews.nvidia.com/news/nvidia-introduces-new-family-of-bluefield-dpus-to-bring-breakthrough-networking-storage-and-security-performance-to-every-data-center>).

²⁶ Ex. 4 (<https://nvidianews.nvidia.com/news/nvidia-extends-data-center-infrastructure-processing-roadmap-with-bluefield-3>).

²⁷ Ex. 11 (<https://blogs.nvidia.com/blog/nvidia-vmware-new-era-enterprise-computing>); see also <https://youtu.be/TuggBO97yYg?si=yxu0LnAN85df-frZ&t=784> (13:04–13:23).

28. Microsoft has also repeatedly praised the significance of Xockets' DPU technology.

29. For example, Microsoft has described building out its Azure infrastructure as "[t]he most important thing we've done over the last four years":

*"The most important thing is what we've done over the last four years [since 2019] to actually build out the core infrastructure on which OpenAI is built: these large models, the training infrastructure — and the [ML/AI] infrastructure doesn't look like regular cloud infrastructure. We had to evolve Azure [with NVIDIA] to have specialized AI infrastructure on which OpenAI is built. . . ."*²⁸

30. In addition, Microsoft has boasted about the benefits the technology brings to Microsoft Azure:

"Together with NVIDIA, we are making the promise of AI real, helping drive new benefits and productivity gains for people and organizations everywhere," said Satya Nadella, chairman and CEO, Microsoft. *"From bringing the GB200 Grace Blackwell processor to Azure, to new integrations between DGX Cloud and Microsoft Fabric, the announcements we are making today will ensure customers have the most comprehensive platforms and tools across every layer of the Copilot stack, from silicon to software, to build their own breakthrough AI capability."*²⁹

31. Microsoft has emphasized the significance of its collaboration with NVIDIA in delivering "state-of-the-art AI capabilities for every enterprise on Microsoft Azure":

"AI is fueling the next wave of automation across enterprises and industrial computing, enabling organizations to do more with less as they navigate economic uncertainties," said Scott Guthrie, executive vice president of the Cloud + AI Group at Microsoft. *"Our collaboration with NVIDIA unlocks the world's most scalable*

²⁸ Ex. 12 (<https://www.theverge.com/23589994/microsoft-ceo-satya-nadella-bing-chatgpt-google-search-ai>).

²⁹ Ex. 13 (<https://news.microsoft.com/2024/03/18/microsoft-and-nvidia-announce-major-integrations-to-accelerate-generative-ai-for-enterprises-everywhere/>).

supercomputer platform, which delivers state-of-the-art AI capabilities for every enterprise on Microsoft Azure.”³⁰

32. Microsoft has boasted that with NVIDIA, it is providing “the most powerful AI supercomputer” in the world to its customers:

“The next wave of computing is being born, between next-generation immersive experiences and advanced foundational AI models, we see the emergence of a new computing platform,” said Satya Nadella, chairman and CEO of Microsoft. *“Together with NVIDIA, we’re focused on both building out services that bridge the digital and physical worlds to automate, simulate and predict every business process, and bringing the most powerful AI supercomputer to customers globally.”*³¹

33. Finally, both NVIDIA and Microsoft publicly promote Microsoft’s use of the accused NVIDIA DPU-enabled systems that copy Xockets’ technology. For example, NVIDIA advertises that “Microsoft Azure and NVIDIA are empowering enterprises to achieve new levels of innovation. With NVIDIA’s *full-stack accelerated computing platform* combined with Microsoft’s global-scale, simplified infrastructure management, enterprises can transform their businesses.”³²

34. NVIDIA also advertises that it is “partnering with Microsoft to accelerate the development and deployment of generative AI across Microsoft Azure, Azure AI services, Microsoft Fabric, and Microsoft 365.”³³

35. NVIDIA also advertises that Microsoft Azure uses NVIDIA’s BlueField DPUs³⁴:

³⁰ Ex. 14 (<https://nvidianews.nvidia.com/news/nvidia-microsoft-accelerate-cloud-enterprise-ai>).

³¹ Ex. 15 (<https://nvidianews.nvidia.com/news/nvidia-and-microsoft-to-bring-the-industrial-metaverse-and-ai-to-hundreds-of-millions-of-enterprise-users-via-azure-cloud>).

³² Ex. 16 (<https://www.nvidia.com/en-us/data-center/dgx-cloud/>).

³³ Ex. 17 (<https://www.nvidia.com/en-us/events/microsoft-build/>).

³⁴ GTC 2023 Keynote with NVIDIA CEO Jensen Huang, <https://www.youtube.com/watch?v=DiGB5uAYKAg&t=1884s> (31:24–31:39).



Microsoft and Nvidia have also entered into numerous agreements relating to the AI technology. For example, they recently announced:³⁵

- Microsoft Azure to Adopt NVIDIA Grace Blackwell Superchip to Accelerate Customer and First-Party AI Offerings
- NVIDIA DGX Cloud's Native Integration with Microsoft Fabric to Streamline Custom AI Model Development with Customer's Own Data
- NVIDIA Omniverse Cloud APIs First on Azure Power Ecosystem of Industrial Design and Simulation Tools
- Microsoft Copilot Enhanced with NVIDIA AI and Accelerated Computing Platforms
- New NVIDIA Generative AI Microservices for Enterprise, Developer and Healthcare Applications Coming to Microsoft Azure AI

Xockets Puts Microsoft and NVIDIA on Notice of Its Patents

Microsoft

³⁵ Ex. 41 (<https://nvidianews.nvidia.com/news/microsoft-nvidia-generative-ai-enterprises>)

36. The following paragraphs (36 to 51) in this section are based on my review of the Xockets' correspondence with Microsoft and NVIDIA.

37. In March 2017, Xockets' Dan Alvarez reached out to Microsoft's Ulrich Homann (Corporate Vice President, Cloud and AI) and Jim Brisimitzis (General Manager, Cloud Developer Relations) with a call for bids. Exhibit 31, at 1. Xockets also provided an overview of Xockets' technology and the fact that Xockets already had a large number of issued patents on the technology:

WHAT DOES XOCKETS DO?


XOCKETS DESIGNS THE XSTREAM APPLIANCE

Public cloud providers, web-scale services companies, and OEMs can directly create new, unique, and powerful **hardware-accelerated services, just by programming software.**

How?

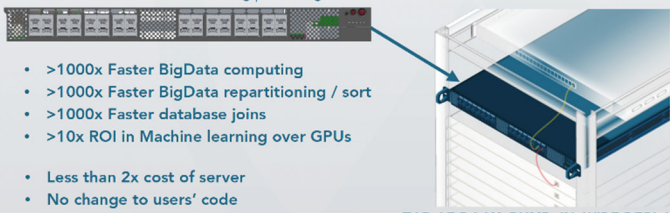
The XStream contains the worlds first physical, streaming processors. Our appliance inserts stream processing into the spine of clusters making the most difficult Machine Learning, batch Map-Reduce, or in-memory streaming analytics applications thousands of times faster, using a fraction of resources.

CONFIDENTIAL AND PROPRIETARY



XSTREAM APPLIANCE


320 Gb/s to 2.2 Tb/s of streaming processing

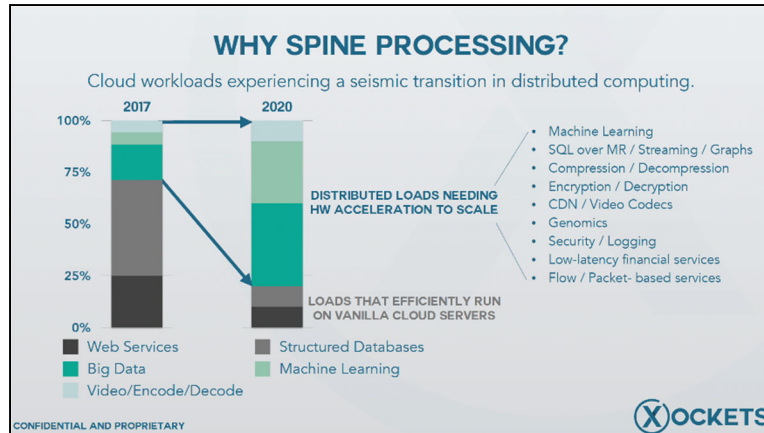


- >1000x Faster BigData computing
- >1000x Faster BigData repartitioning / sort
- >1000x Faster database joins
- >10x ROI in Machine learning over GPUs
- Less than 2x cost of server
- No change to users' code
- Available for Hadoop and Spark demonstrations today

TOP OF RACK, BUMP-IN-WIRE DEPLOYMENT
XStream inserts reconfigurable, streaming processors into the switching spine of clusters

CONFIDENTIAL AND PROPRIETARY





IP PORTFOLIO

BROAD COVERAGE OF DISTRIBUTED ACCELERATION

- Xockets IP portfolio is **fundamental to acceleration fabrics in clusters**. Over 50 patents, with over a dozen foreign-filed and granted, the Xockets patent portfolio was built for **strategic licensing (for ODMs) and provide offensive/defensive IP in a highly competitive market**.
- The patents broadly cover the seminal architectures of cluster acceleration using **network side acceleration, and IO cache acceleration**.
- The technology developed will produce hundreds more patents in the future that broadly apply to today's and future markets.

CONFIDENTIAL AND PROPRIETARY

XOCKETS

A copy of this presentation is attached as Exhibit 18.

38. In response, Mr. Homann responded that the “concept resonates and the team would like to understand in more depth.” Ex. 31, at 1. Mr. Homann directed Xockets to interface with Saurabh Kulkarni (Director of Engineering, Cloud and AI System Technologies) and Kushagra Vaid (VP and Distinguished Engineer, Azure Infrastructure). Ultimately, Dr. Dalal had a discussion with Mr. Kulkarni and Tanj Bennett (Partner SDE) on March 22, 2017, so they could “get a technical overview of key Xockets technologies in the hardware acceleration space.” Ex. 32, at 2. Thereafter, Mr. Kulkarni informed Dr. Dalal that he was reaching out to folks from Microsoft’s “big data and machine learning teams” in order to make an introduction. *Id.* at 1.

NVIDIA

39. NVIDIA has had actual notice of Xockets' New Cloud Processor and New Cloud Fabric Patents at least since February 2022.

40. On January 27, 2022, Xockets founder, Dr. Dalal, communicated with NVIDIA's Brad Genereaux (Global Lead, Healthcare Alliances) and asked for an introduction to "NVIDIA legal IP Counsel" in order to discuss "some very strategic IP" that "NVIDIA would be interested in acquiring." Ex. 40. After making an internal inquiry, Mr. Genereaux ultimately connected Dr. Dalal with Gady Rosenfeld on February 4, 2022. Ex. 30, at 3.

41. Mr. Rosenfeld was "leading the DPU segment in the NVIDIA field organization" at that time. Ex. 30, at 3. Indeed, Mr. Rosenfeld's LinkedIn profile reflects that he has been NVIDIA's Vice President, DPU Business since July 2021 and remains in that role today.³⁶

42. After connecting, Dr. Dalal emailed Mr. Rosenfeld sample claim charts and a list of Xockets' then-current patent list covering breakthrough DPU technologies essential to AI, which included the New Cloud Processor Patents (the '209, '924, and '350 Patents) and the New Cloud Fabric Patents (the '297, '161, '092, and '640 Patents). Ex. 30, at 1.

43. To date, NVIDIA and Microsoft have refused to negotiate individually with Xockets for a patent license or acquisition. In 2024, Xockets again reached out to NVIDIA and Microsoft to commence negotiations, but neither would engaged. The history of these discussions are set forth in the Declaration of Brian Hinman (attached as Exhibit 34) and the Declaration of Charles Fish (attached as Exhibit 35).

RPX's Description of Its Business Model

44. RPX was founded in 2008 and has more than 450 members. It has issued numerous

³⁶ Ex. 33 (<https://www.linkedin.com/in/gadyrosenfeld/>).

public statements describing its business model. These include the following³⁷:

RPX Corporation brings companies together from throughout the world to solve patent risks that they face in common. Our conviction is that solving such problems once for many companies can achieve a faster, better, and less expensive resolution than might otherwise be achieved by each company acting alone. To this end, we offer a platform that includes defensive buying of patent rights, acquisition syndication, patent intelligence, insurance services, and advisory services.

Our pioneering approach combines principal capital, deep patent expertise, and client contributions to generate enhanced patent buying power. By efficiently acquiring rights to problematic patents, we help to mitigate and manage the risk of potential patent assertions for our growing client network.

45. RPX previously touted on its website (language that has since been removed) that “[i]n effect, RPX can buy ‘wholesale’ on behalf of our client network, while our clients otherwise would pay ‘retail’ if transacting on their own.” A copy of this website is attached as Exhibit 20.

46. The RPX website also previously advertised that “RPX is often able to achieve ‘wholesale’ pricing terms, where we can acquire rights for our members at significantly reduced cost relative to what the NPE might charge an individual company on its own. RPX believes we have saved our members tens of millions of dollars through these wholesale-priced transactions.” A copy of this website is attached as Exhibit 21. RPX’s most recent 10-K filing with the SEC in 2018 explains its mission of interjecting itself as the “essential intermediary” between patent owners and RPX’s members:

Our mission is to reduce risk and cost for corporate legal departments through data-driven decision-making, technology, and market-based solutions. A significant part of that mission is to transform the patent market by establishing RPX as the essential intermediary between patent owners and operating companies and by providing complementary technology-focused discovery services. Our strategy includes the following:

Ex. 22, at 6.

47. RPX’s co-founder and former CEO, John Amster, has publicized RPX’s mission, stating that “[w]e think there can be a clearinghouse in this market that can be really quite big and

³⁷ Ex. 19 (<https://www.rpxcorp.com/about/>).

efficient. If every company just decided, ‘We’re going to have a line item in our budget for patents and patent risks, and that line item is going to be the RPX rate card’—i.e. RPX’s subscription rate[.]”

48. Indeed, RPX’s website (which refers to the “RPX Network” as the “world’s leading defensive patent acquisition network”) touts how RPX’s application of “capital to acquire patents rights” leads to “far less cost” for its members, that “[t]here’s safety in numbers” and “huge cost savings, too.”³⁸

49. RPX’s website also explains how it collaborates with its members and non-members to create anti-competitive buyers’ cartels, what it euphemistically calls “syndicated licensing transactions”³⁹:

In addition to our core patent acquisition service, RPX also facilitates large-scale syndicated licensing transactions that can include non-members and members (who make contributions beyond their regular subscription fees).

50. RPX previously highlighted the benefits of these syndicated transactions for its clients on its website, in language that has since been removed, stating that “[o]ur clients see distinct advantages of syndicated purchasing through RPX, as we are uniquely situated to structure transactions that are ultimately less costly and deliver more value to participating clients than if any attempted individual licensing or unilateral purchasing of the portfolios.” Ex. 20.

51. RPX has openly acknowledged in its SEC filings that its practices may be illegal and violate competition and antitrust laws, admitting that “[i]t is possible that courts or other

³⁸ Ex. 23 (www.rpxcorp.com/solutions/rpx-network).

³⁹ Ex. 24 (<http://ir.rpxcorp.com>).

governmental authorities will interpret existing laws regulating [] competition and antitrust practices [] in a manner that is inconsistent with our business practices.”⁴⁰

Investigations of NVIDIA and Microsoft for Antitrust Violations

52. NVIDIA is currently under investigation in multiple jurisdictions, including the United States, for anticompetitive conduct, including price fixing. For example, the Department of Justice recently initiated “two separate probes into Nvidia regarding antitrust concerns.”⁴¹ It faces further scrutiny from antitrust regulators in Europe and Britain. In a July 30, 2024 Letter to the DOJ, several public interest groups noted that “Nvidia’s anti-competitive practices have raised grave concerns about price-fixing and bundling — concerns that have already led to French antitrust enforcement action and scrutiny by regulators in the European Union and the United Kingdom.” This letter is attached as Exhibit 27.

53. France’s national competition regulator, the Autorite de la concurrence, recently issued report outlining “[t]he risk of abuse by chip providers,” specifically NVIDIA, within the generative AI industry. These include “price fixing, production restrictions, unfair contractual conditions and discriminatory behavior.”⁴²

54. Microsoft is also under anti-trust investigation because of its de facto control over the dominant large language artificial intelligence models in the world.⁴³

⁴⁰ Ex. 25, RPX Corporation, S.E.C. Registration Statement (Form S-1), at 17 (Jan. 21, 2011), *available at* <https://www.sec.gov/Archives/edgar/data/1509432/000119312511012087/ds1.htm>.

⁴¹ Ex. 26 (<https://www.msn.com/en-us/news/technology/nvidia-faces-two-doj-antitrust-probes-over-market-dominance/ar-BB1r4Ui7?ocid=BingNewsVerp>).

⁴² Ex. 28 (<https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive>).

⁴³ Ex. 29 (<https://www.nytimes.com/2024/06/05/technology/nvidia-microsoft-openai-antitrust-doj-ftc.html>).

Claim Charts

55. Attached as Exhibits 37-40 are exemplary charts showing that NVIDIA's about to launch Blackwell architecture that Microsoft is integrating into its GPU-enable AI platform uses Xockets' patented technology.

I declare under penalty of perjury and the laws of the United States of America that the foregoing is true.

Executed on September 5, 2024

By: /s/ Jason Sheasby

Jason Sheasby